

Information Characteristics of Social Activity

Youri Arzumanyan¹ & Arian Zakharov¹

¹ Department of Data Economics, The Bonch-Bruevich St Petersburg State University of Telecommunications (SPbSUT), St. Petersburg, Russia.

Correspondence: Arian Zakharov, Department of Data Economics, The Bonch-Bruevich St Petersburg State University of Telecommunications (SPbSUT), St. Petersburg, Russia.

Received: September 21, 2025 Accepted: October 20, 2025 Published: November 11, 2025

doi:10.65343/erd.v1i2.42 URL: https://doi.org/10.65343/erd.v1i2.42

Abstract

Assuming that the types of an individual's (agent's) social activity form a finite countable set, and employing the characteristics of digital traces, the model of agents' actions is considered as a Markov information source whose diversity is described by the sum of constructive and destructive entropies. A coefficient of constructive development is introduced as the ratio of constructive entropy to the entropy of the overall diversity of social activity. The rate of an agent's social transformation is defined as the ratio of mutual information about the transformation to the transformation time interval. The maximum value of this rate, given a specified stability coefficient, defines the constructive transformational capacity of the social environment. Obstacles to the practical validation of the proposed characteristics are discussed.

Keywords: social activity and development, Markov model of agents, information theory, development stability, rate of social transformation, transformational capacity of the social environment

1. Introduction

Starting with the fundamental monographs (Ashby, 1956; Theil, 1967), and up to the present time, many researchers have used numerical characteristics of information theory to analyze economic and social phenomena, as for example in (Hao, Gao, Lin, Wu, & Shang, 2024; Mucciardi, & Benassi, 2025), where information entropy or characteristics based on it are applied to the analysis of social processes. Nevertheless, the authors believe that the potential capabilities of information theory for solving problems in sociology are far from exhausted and are acquiring even greater significance in the modern post-industrial world. The reason for this conclusion is the features of development noted in (Arzumanyan, 2021; 2022; 2023) for both individual persons (agents) and social groups. These features can include the following:

- 1). Modern information technologies, traffic control methods, and geolocation allow for obtaining objective data on the social activity (microstate) of individual individuals (agents), and the nature of personal activity has an increasingly greater influence on the activity of social groups and society as a whole (Smart World concept).
- 2). The totality of data on the microstates of agents allows for a sufficiently complete characterization of their social activity. In other words, the projection of agents' physical actions onto the information space sufficiently fully reflects social phenomena, which can be judged by information traces, thereby defining a morphism between the agents' physical world and the virtual information environment.
- 3). All identified types of agents' microstates can be enumerated and numbered, forming a countable set of microstates.

The authors attempted to use these features by applying the results of information theory to provide an objective description of the processes of development of social activity.

2. Materials and Methods

Based on the fact that in the modern world, digital footprints sufficiently fully characterize the social activity of an individual (agent), the actions of agents can be considered as a Markov source of information, the diversity of which is described by Shannon entropy. The social activity of the agent manifests itself in the social environment with its own rules and restrictions. Leaving aside moral and ethical evaluations of the meaning of these prohibitions, the entire set of types of agent activities can be

formally divided into two disjoint subsets such that one includes all permissible social manifestations, and the other only prohibited ones. Such a division allows representing the entropy as two components (constructive and destructive entropies) and defining the coefficient of development stability as the ratio of the constructive component to the total entropy.

Interpreting the development (transformation) of the agent as a process of transmission over a communication channel, the role of which is played by the social environment, one can define the speed of social transformation of the agent as the ratio of mutual information about the transformation to the transformation time interval. The maximum value of this speed at a given stability coefficient determines the constructive transformational capacity of the social environment.

For the translation of the article text into English, Grok 4 and ChatGPT 5 was used.

3. Information Characteristics of Social Activity

3.1 Markov Model of Agent Space

We shall use the term *agent* to denote an individual exhibiting social activity within a social environment (a social group or society). The fact of demonstrating social activity is interpreted as a microstate of the agent. All possible types of agent microstates can be defined and enumerated in such a way that, at any arbitrary moment in time, the agent is in exactly one type of microstate.

Let us denote by **m** the set of agent A's microstate types:

$$\mathbf{m} = \{S_1, S_2, ..., S_i, ..., S_M\},\$$

where S_i – is a microstate of the i- th type, and

M - is the total number of microstate types of agent A.

The current microstate of agent A may significantly depend on its previous states. To take this into account, while avoiding cumbersome computations, we shall assume that the sequence of agent microstates forms a first-order Markov chain, such that the presence of the agent in any microstate $S_i \in \mathfrak{m}$ depends only on its immediately preceding microstate.

Based on the observed microstates of the agent, for each element $S_i \in \mathbf{m}$ one can estimate the conditional probability

$$p(S_i/S_j) = p_{ij},$$

which represents the probability that, at an arbitrary moment in time, the agent is in microstate S_i , given that its previous microstate was S_j . The sum of these probabilities yields the unconditional probability of the agent being in a given microstate

$$p(S_i) = p_i = \sum_{j=1}^{M} p_{ij},$$

where

 p_i – is the unconditional probability of the agent being in microstate S_i ,

 p_{ij} is the conditional probability of the agent being in microstate of type i,

given that its previous microstate was of type j,

M – is the total number of microstate types of agent A.

The set of probabilities $\{p_{ij}\}$, together with the elements of \mathbf{m} , form the ensemble \mathbf{s} of messages concerning the social activity of agent \mathbf{A}

$$\mathbf{\mathfrak{S}} = \{ (S_1, p_{11}, p_{12} ..., p_{1M}), ..., (S_i, p_{i1}, ..., p_{ij}, ..., p_{iM}), ..., p_{MM}) \}, \tag{1}$$

Where

 \mathfrak{S} – is the ensemble of messages on the social activity of agent A,

 S_i – is a microstate of the *i*-th type,

 p_{ij} – is the conditional probability of agent **A** being in microstate of type i,

given that the previous microstate was of type j,

M - is the number of agent A's microstate types.

As a measure of the diversity of agent A's microstates (Ashby, 1956), one may use the entropy **H**(**5**) of the ensemble **5** (MacKay, 2003)

$$\mathbf{H}(\mathbf{S}) = \sum_{i=1}^{M} p_i \, \mathbf{H}(\mathbf{S} / S_i), \tag{2}$$

where

$$p_{i} = \sum_{j=1}^{M} p_{ij},$$

is the unconditional probability of agent A being in microstate S_i , and

$$\mathbf{H}(\mathbf{S} / S_i) = \sum_{j=1}^{M} p_{ij} \log_2(1 / p_{ij}),$$

is the conditional entropy for microstates of the i-th type.

Here,

 p_{ij} is the conditional probability of agent **A** being in microstate of type i, given that the previous microstate was of type j,

M is the total number of microstate types of agent A.

3.2 Development Characteristics

The social activity of an agent is meaningful to analyze only when this activity is manifested within a social environment governed by its own rules and constraints. Leaving aside moral and ethical judgments concerning the content of such restrictions, the entire set \mathbf{m} of agent microstate types can formally be divided into two disjoint subsets: one containing all permissible manifestations of social activity, denoted \mathbf{m}^+ , and the other containing only prohibited ones, denoted \mathbf{m}^- .

This partition makes it possible to express the entropy $\mathbf{H}(\mathbf{S})$ as the sum of two components:

$$H(S) = H(S)^{+} + H(S)^{-}$$

where

$$\mathbf{H}(\mathbf{S})^{+} = \sum p_i \mathbf{H}(\mathbf{S} / S_i),$$

$$\mathrm{index}(S_i \in \mathbf{m}^{+})$$

is the entropy of permissible types of activity, and

$$\mathbf{H}(\mathfrak{S})^{-} = \sum_{i \in \mathcal{S}_i} p_i \mathbf{H}(\mathfrak{S}_i / S_i)$$

$$index(S_i \in \mathfrak{m}^{-})$$

is the entropy of prohibited types of activity.

Here, index($S_i \in \mathbf{x}$) indicates that the summation is taken over the indices of microstates S_i belonging to the set of types \mathbf{x} (either \mathbf{m}^+ or \mathbf{m}^-).

It is important to emphasize that representing entropy as the sum of two components does not imply that the ensemble **5** can be divided into two independent ensembles. On the contrary, in the calculation of each component, microstates of all types are taken into account; only the summation procedure with respect to the first index is partitioned into two parts.

The term $\mathbf{H}(\mathbf{S})^+$ characterizes the diversity of activities supported by the social environment and contributing to the agent's development as a member of that environment. Conversely, the term $\mathbf{H}(\mathbf{S})^-$ defines the diversity of activities that pose a potential threat to the agent's social significance and have a destructive character.

Thus, the balance of an agent's activity—whether it tends toward constructive development or destructive influence within the social environment—can be assessed using the ratio of the constructive diversity $\mathbf{H}(\mathbf{S})^+$ to the total diversity $\mathbf{H}(\mathbf{S})$. We shall refer to this ratio as the **coefficient of social stability**:

$$\alpha(\mathfrak{S}) = H(\mathfrak{S})^+ / H(\mathfrak{S})$$

where $a(\mathfrak{S})$ denotes the coefficient of social stability of agent A (ensemble \mathfrak{S}).

It is straightforward to observe that $0 \le \alpha(\clubsuit) \le 1$. The boundary values are attained either in the case of purely constructive diversity of the agent's microstates, when

 $\mathbf{H}(\mathbf{S})^+ \equiv \mathbf{H}(\mathbf{S}) \ (\mathbf{H}(\mathbf{S})^- = 0)$ and $\alpha(\mathbf{S}) = 1$, or in the case of purely destructive diversity, when $\mathbf{H}(\mathbf{S})^- \equiv$

$$\mathbf{H}(\mathbf{S}) \ (\mathbf{H}(\mathbf{S})^{+} = 0) \text{ and } \alpha(\mathbf{S}) = 0.$$

To illustrate the meaning of the coefficient $\alpha(\mathbf{S})$, let us consider an example of ensemble \mathbf{S} with four (M=4) microstate types of agent A (Table 1).

Table 1. Ensemble \$\sigma\$ of microstates of agent A

			Previous microstate					
Current microstate	m		m ⁺			m-	Unconditional probability of the microstate	
			S ₁	S ₂	S ₃	S ₄	of the inicrostate	
sleep, rest at home		S ₁	$p_{11} = 0.1$	$p_{12} = 0.1$	$p_{13} = 0.1$	$p_{14} = 0.1$	$p_1 = p_{11} + p_{12} + p_{13} + p_{14} = 0.4$	
transfer to/from home	m+	S ₂	$p_{21} = 0.02$	$p_{22} = 0.005$	$p_{23} = 0.02$	$p_{24} = 0.005$	$p_2 = p_{21} + p_{22} + p_{23} + p_{24} = 0.05$	
work	_	S ₃	$p_{31} = 0.05$	$p_{32} = 0.1$	$p_{33} = 0.25$	$p_{34} = 0.1$	$p_3 = p_{31} + p_{32} + p_{33} + p_{34} = 0.5$	
prohibited activity	m-	S ₄	$p_{41} = 0.02$	$p_{42} = 0.005$	$p_{43} = 0.02$	$p_{44} = 0.005$	$p_4 = p_{41} + p_{42} + p_{43} + p_{44} = 0.05$	
							$p_1 + p_2 + p_3 + p_4 = 1$	

Table 2 presents the results of calculating the conditional entropies, the entropy $\mathbf{H}(\mathbf{S})$, and the stability coefficient of ensemble \mathbf{S} .

Table 2. Entropies and stability coefficient of ensemble \$\simes\$

	Conditiona	Entropy	Stability		
H(\$\sqrt{1}\)	$\mathbf{H}(\mathbf{S} / S_2)$	H(\$\sigma / S_3)	H(\$\sigma\) / S4)	H(\$)	α (§)
1.33	0.30	1.38	0.30	1.25	0.988

It is not difficult to observe that the results of monitoring the agent's stability coefficient can be employed as a criterion or as an objective function in problems of personal optimization of development, as well as in prospective planning of group strategies.

If we assume that the sequence of agent microstates forms an n-th order Markov chain, then expression (1) for the ensemble S of messages on the social activity of agent A takes the form

$$\mathfrak{S} = \{(S_1, \mathbb{P}_1), \ldots, (S_i, \mathbb{P}_i), \ldots, (S_M, \mathbb{P}_M)\},$$

where

 S_i – is a microstate of the *i*-th type,

 $\mathbb{P}_i = (p_{i11...1}, \dots, p_{ij_1j_2...j_n}, \dots, p_{iMM...M})$ is an (n+1)-dimensional vector of conditional probabilities of agent A being in a microstate of type i, given the previous n microstates of all possible types, M is the total number of agent A's microstate types.

Formula (2) for computing the entropy **H(\$)** of ensemble **\$** can be rewritten as

$$\mathbf{H}(\mathbf{S}) = \sum_{i=1}^{M} p_i \, \mathbf{H}(\mathbf{S} / S_i),$$

where

$$p_i = \sum_{j_1, j_2, \dots, j_n}^{\mathbf{M}} \mathbb{P}_i$$

is the unconditional probability of agent A being in microstate S_i,

$$\mathbf{H}(\mathfrak{S}/S_i) = \sum_{j_i,j_2,...,j_n=1}^{M} \mathbb{P}_i \log_2(1/\mathbb{P}_i)$$

is the conditional entropy for the microstates of type i,

$$\mathbb{P}_i = (p_{i11..1}, \dots, p_{ij_1j_2..j_n}, \dots, p_{iMM..M})$$

is an (n+1)-dimensional vector of conditional probabilities of agent A being in a microstate of type i, given the previous n microstates of all possible types, M is the total number of agent A's microstate types. All other expressions remain unchanged.

3.3 Characteristics of Development Dynamics

Of particular interest are the characteristics of development dynamics, among which the most important is the rate of change in the diversity of the agent's microstate types, i.e., the rate of social development.

Let us denote by \mathbf{n} the set of agent microstate types after a time interval Δt :

$$\mathbf{n} = \{Y_1, Y_2, ..., Y_i, ..., Y_N\},\$$

where

 Y_i is a microstate of the *i*-th type,

N is the number of agent A's microstate types after time interval Δt .

In the set **n**, new microstate types absent from **m** may emerge, while some previously existing types may disappear. In other words, both the number and the composition of microstate types at the initial moment and after the interval Δt need not be the same (the sets \mathbf{n} and \mathbf{m} do not necessarily coincide).

In classical information theory, a similar situation arises when analyzing communication systems in which the recipient's ensemble contains one new message, referred to as an erasure. For the social environment, however, there may be several such "erasures."

Assuming that, at the end of the interval Δt , the Markov model of the agent's microstates still preserves its n-th order, the ensemble \mathfrak{P} of agent microstates at the end of the interval can be written as

$$\mathbf{P} = \{(Y_1, \mathbb{Q}_1) ..., (Y_i, \mathbb{Q}_i) ..., (Y_N, \mathbb{Q}_N)\},\$$

where

 Y_i is a microstate of the *i*-th type at the end of the time interval Δt ,

$$\mathbb{Q}_{i} = (q_{i11..1}, \dots, q_{ij_{1}j_{2}..j_{n}}, \dots, q_{iNN..N})$$

is an(n+1)-dimensional vector of conditional probabilities of agent A being, after interval Δt , in a microstate of type i, given the previous n microstates of all possible types, N is the number of agent A's microstate types after the time interval Δt .

The value **H**(♠) of the final diversity of agent microstate types can be defined as

$$\mathbf{H}(\mathbf{H}) = \sum_{i=1}^{N} q_i \, \mathbf{H}(\mathbf{H} / \mathbf{Y}_i),$$

where

$$q_i = \sum_{j_1, j_2, \dots, j_n = 1}^{N} \mathbb{Q}i,$$

is the unconditional probability of a microstate Y_i of type i, and

$$\mathbf{H}(\mathbf{P}/Y_i) = \sum_{j_1,j_2,...,j_{n_i}=1}^{N} \mathbb{Q}_i \log_2(1/\mathbb{Q}_i)$$
 is the conditional entropy for microstates of type i ,

$$\mathbb{Q}_{i} = (q_{i11..1}, \dots, q_{ij_{1}j_{2}..j_{n}}, \dots, q_{iNN..N})$$

is an (n+1 dimensional vector of conditional probabilities of agent A being in a microstate of type i after interval Δt , given the previous n microstates of all possible types, N is the number of agent A's microstate types after the time interval Δt .

From the perspective of information theory, the process of transforming ensemble \$\mathbb{G}\$ into ensemble \$\mathbb{P}\$ can be interpreted as the transmission of messages through a communication channel, which is represented by the social environment surrounding the agent. This projection makes it possible to characterize the social development of the agent as

$$I = H(\mathfrak{P}) - H(\mathfrak{P} / \mathfrak{S}),$$

where

I is the amount of mutual information concerning the agent's social transformation,

H(♠) is the final diversity of the agent's microstate types,

H(♠ /♠) is the conditional final diversity given the initial ensemble ♠.

In a natural way, the rate of social development of agent A over the interval Δt can be defined as

$$\mathbf{V} = \mathbf{I} / \Delta t$$
.

where

V is the rate of social development of agent A,

I is the amount of mutual information concerning the social transformation of A,

 Δt is the duration of the transformation.

As in information theory, the conditional entropy $\mathbf{H}(\mathfrak{Y})$ / \mathfrak{S}) characterizes the influence of the social environment (the "communication channel") on agent \mathbf{A} . To compute the value of $\mathbf{H}(\mathfrak{Y})$ / \mathfrak{S}), the probabilities r_{ij} must be known, representing the probability of the transformation, during interval Δt , of a microstate of type i from ensemble \mathfrak{S} into a microstate of type j in ensemble \mathfrak{Y} . In this case,

$$\mathbf{H}(\mathbf{P} / \mathbf{S}) = \sum_{j=1}^{M} r_i \, \mathbf{H}(Y_i / S_j),$$

where

$$r_i = \sum_{j=1}^{M} r_{ij},$$

is the unconditional probability of transformation into microstate Y_i after Δt , and

$$\mathbf{H}(Y_i / S_j) = \sum_{i=1}^{N} r_{ij} \log_2(1/r_{ij}),$$

is the conditional entropy of transformation from S_j into Y_i .

Here,

 r_{ij} is the conditional probability of transformation into microstate Y_i from S_i ,

M is the number of microstate types in the initial ensemble \$5,

N is the number of microstate types in the final ensemble **2.**

The interval Δt is assumed to be sufficiently large so that it is adequate to rely solely on these probabilities.

Just as the stability coefficient $\alpha(\$)$, the rate of social development V can be used as a criterion or an objective function in problems of analysis and synthesis of the social environment. In addition, the maximum value of V, obtained by enumerating all possible ensembles \$ (that is, all agents of a social group or of society as a whole), determines the **group transformational capacity** (an analogue of channel capacity in classical information theory)

$$T = \max_{\mathbf{S}} V$$
,

where

T is the transformational capacity of the social environment of agent A,

V is the rate of social development of the members of the social group (agents),

\$ are the ensembles of agents' social activity.

Since, in the enumeration used to determine the transformational capacity, ensembles of all agents are taken into account, including those with a low stability coefficient, it appears reasonable to restrict the enumeration to only those ensembles whose stability coefficient exceeds a given threshold. This characteristic of the social environment may be referred to as the **constructive transformational capacity** T_{α}

$$T_{\alpha} = \max V$$
,
 \mathfrak{S} , $\alpha(\mathfrak{S}) \geq \alpha_0$

where

 T_{α} is the constructive transformational capacity of the social environment,

V is the rate of social development of the members of the social group (agents),

s are the ensembles of agents' social activity,

 $\alpha(\mathbf{S})$ re the stability coefficients of the agents,

 α_0 is the prescribed threshold value of the coefficient $\alpha(\mathfrak{S})$.

As an example, let us consider the situation where all unconditional probabilities of the microstates in ensemble \mathfrak{S} at the initial moment are equal, and therefore each is equal to 1/M ($\forall i \in \{1,...M\}$ $p_i = 1/M$). At the end of the interval Δt , ensemble \mathfrak{S} contains new types of microstates ($N \geq M$). Let the probability that any microstate from ensemble \mathfrak{S} is lost during the transformation process, without being converted into any of the new microstates of ensemble \mathfrak{S} , be denoted by p, and the probability that such a conversion does occur be denoted by r.

In the absence of restrictions on the constructive nature of microstate types ($\alpha_0 = 0$), the described situation corresponds, in information theory, to transmission over an M-ary symmetric memoryless channel with equiprobable (N-M) distinct erasure symbols. Consequently, the transformational capacity of the environment is equal to the capacity of such a channel.

As can be readily shown, in this case

$$\mathbf{T} = (1 - p - r) \log_2(1 - p - r) + p \log_2[p / (M - 1)] - (1 - r) \log_2[(1 - r) / M],$$

With $\mathbf{T}_{min} = 0$ when $M p = (M - 1) (1 - r)$.

If we assume that there is no loss of microstate types $p \equiv 0$, then the expression for the transformational capacity T simplifies to

$$\mathbf{T} = (1 - r) \log_2 \mathbf{M}$$

It is easy to observe that this last expression leads to the obvious conclusion that the transformational capacity increases with the growth of the diversity of agents' microstates.

4. Discussion

The examples presented of characteristics obtained based on information theory are of a debatable nature since they require verification on real statistics, access to which is limited by laws on the protection of personal data. However, the unconditional adequacy of information theory and its objective nature determine the potential prospects for applying the presented characteristics to the analysis of the current process of social development of individuals and social groups and may be useful for demonstrating the capabilities of information theory in master's program courses in informatics and sociology.

References

- Arzumanyan, Y. (2021). Social Environment Engineering. X INTERNATIONAL CONFERENCE ON ADVANCED INFOTELECOMMUNICATIONS ICAIT. *Conference Proceedings, vol. 3*, pp. 325-330. Retrieved from https://www.sut.ru/doci/nauka/1AEA/APINO/10-APINO-2021.%20T.3.pdf
- Arzumanyan, Y. (2022). Entropy of the Social Environment. XI INTERNATIONAL CONFERENCE ON ADVANCED INFOTELECOMMUNICATIONS ICAIT. *Conference Proceedings, vol. 3*, pp. 552-556. Retrieved from https://www.sut.ru/doci/nauka/1AEA/APINO/11-APINO-2022.%20%D0%A2.3.pdf
- Arzumanyan, Y. (2023). Development and Sustainability of Postinformation Society (Smart World). XII INTERNATIONAL CONFERENCE ON ADVANCED INFOTELECOMMUNICATIONS ICAIT. Conference Proceedings, vol. 4, pp. 9-15. Retrieved from https://apino.sut.ru/@/file/Ua83P4CoSaFyGo14
- Ashby, W. R. (1956). *An Introduction to Cybernetics*. John Wiley & Sons: New York. Retrieved 5 August 2025, from https://archive.org/download/introductiontocy00ashb/introductiontocy00ashb.pdf
- Hao, F., Gao, J., Lin, Y. G., Wu, Y. L., & Shang, J. X. (2024). *Concept Stability Entropy: A Novel Group Cohesion Measure in Social Networks*. IEEE Trans. Emerging Topics in Computing. Retrieved from https://openreview.net/forum?id=pzgyS19PTt
- MacKay, A. (2003). Information Theory, Inference, and Learning Algorithms. Cambridge University

Press. Retrieved from https://www.inference.org.uk/itprnn/book.pdf

Mucciardi, M., & Benassi, F. (2025). Spatializing Shannon entropy: a Gaussian Kernel approach to studying the territorial distribution of selected foreign population groups in Italy. *Quality & Quantity*, 59(4), 995-1018. https://doi.org/10.1007/s11135-025-02241-4

Theil, H. (1967). *Economics and Information Theory*. North-Holland Publishing Co. & Rand McNally & Company: Amsterdam & Chicago. Retrieved 5 August 2025, https://archive.org/details/economicsinforma0000thei

Copyrights

The journal retains exclusive first publication rights to this original, unpublished manuscript, which remains the authors' intellectual property. As an open-access journal, it permits non-commercial sharing with attribution under the Creative Commons Attribution 4.0 International License (CC BY 4.0), complying with COPE (Committee on Publication Ethics) guidelines. All content is archived in public repositories to ensure transparency and accessibility.